



UPPSALA
UNIVERSITET

Molecular Biotechnology Programme

Uppsala University School of Engineering

UPTEC X 05 022	Date of issue 2005-05	
Author	Mikael Wallman	
Title (English)	Performance estimation in small sample parameter and classifier selection	
Title (Swedish)		
Abstract	<p>In this work it was shown that improper use of performance estimation on classifiers optimized by cross validation induces a positive bias in the error estimate, especially for small sample number data sets typically used in biological applications. The mechanism behind the bias was investigated, and a remedy was suggested in the form of a double cross-validation loop. However, optimization procedures in conjunction with requirements for unbiased error estimates give less reliable performance estimation. This problem was addressed with Bayesian inference and consequences of the decreased reliability were investigated.</p>	
Keywords	performance estimation, bias, cross-validation, Bayesian inference.	
Supervisors	Mats Gustafsson Department of engineering sciences, Uppsala University	
Scientific reviewer	Jan Komorowski LCB, Uppsala University	
Project name	Sponsors	
Language	Security	
English		
ISSN 1401-2138	Classification	
Supplementary bibliographical information	Pages	
	25	
Biology Education Centre Box 592 S-75124 Uppsala	Biomedical Center Tel +46 (0)18 4710000	Husargatan 3 Uppsala Fax +46 (0)18 555217

